



Universiteit
Leiden
The Netherlands

The Impact of Leaders on Electoral Choice in the Netherlands -Revisited

Aarts, K.

Citation

Aarts, K. (2001). The Impact of Leaders on Electoral Choice in the Netherlands -Revisited. *Acta Politica*, 36: 2001(4), 380-401. Retrieved from <https://hdl.handle.net/1887/3450767>

Version: Publisher's Version

License: [Leiden University Non-exclusive license](#)

Downloaded from: <https://hdl.handle.net/1887/3450767>

Note: To cite this publication please use the final published version (if applicable).

The Impact of Leaders on Electoral Choice in the Netherlands - Revisited¹

Kees Aarts

University of Twente

Abstract

In a recent contribution to *Acta Politica*, Van Wijnen (2000) argued that party leaders have become more important for the voting in the Netherlands. This finding is criticized on the ground of the research strategy that was employed. The trade-off between unreliability (as a result of multicollinearity) and bias (as a result of omitted variables) has been ignored. When comparing the consequences of the alternative design choices, it appears that allowing multicollinearity has less serious consequences than omitting important variables. Secondary analysis using a newly specified model shows that there is no increase in the impact of political leaders on voting in the Netherlands.

1 Introduction

The impact of political personalities on the voters' electoral choices has long been a favourite topic in political journalism as well as in popular political debates. Party leaders supposedly can win and lose large numbers of seats for their parties, not just by surfing the waves of economic growth (or failing to do so) but also by their personal appearances, their voices and clothes. The Dutch Christian Democrats' important electoral success in 1986 will for many always be connected with the personal performance of Ruud Lubbers, just as Elco Brinkman has been, and will continue to be, blamed for their disastrous result in 1994. Other examples are readily found.

But showing convincingly that voters actually base their electoral choice to a large, or growing, extent on an assessment of the candidates is quite a different matter. Assessing the impact of political leaders on the vote is notoriously difficult from a methodological point of view.² As part of the electoral options, political leaders are almost inextricably tied up with their parties. As explanations of electoral choice, the evaluations of political leaders usually go up and down with the evaluations of their parties, and both types of evaluations draw on a common set of characteristics of the individual voter.

In a recent contribution, Van Wijnen (2000) attempted to shed light on the question whether candidates have become more important for vote choice in

recent Dutch parliamentary elections. Van Wijnen's contribution fills a gap in the literature on candidates and voting behaviour in the Netherlands. He is aware of some of the research problems just referred to (cf. his section 4.2). He acknowledges that the net impact of candidates on vote choice can only be assessed when a number of other explanations of voting behaviour are taken into account, such as the voter's social background and his or her political opinions. Moreover, Van Wijnen's application of the conditional logit model in estimating the effects of candidates reflects the state-of-the-art in modelling political choice, and is a major improvement over much of the older literature on the topic.

However, Van Wijnen's analysis also suffers from a serious shortcoming that appears to upset some of his central results. He fails to take into account the relationship between the general evaluation of a political party and the corresponding evaluation of the leader of that party. These two objects of evaluation are both important for the decision which electoral choice to make. Focusing on only one of the two, i.e., the leader evaluation, may lead to serious mistakes in estimates of the impact of candidates on vote choice, as I will show in this comment.

The research problem with which I will deal, is interesting beyond the scope of this particular application, since it represents a general dilemma in formulating and testing empirical models in political science. This is the dilemma between including correlated variables versus excluding relevant variables. Alternatively, it is known as the dilemma between multicollinearity and specification error, or between bias and efficiency.

In the next section, Van Wijnen's research strategy is summarized. Thereafter, the general problem associated with this strategy is addressed, which leads to the formulation of two design options. When the design options have been formulated, the data are re-analysed in order to show that Van Wijnen chose the wrong option.

2 Van Wijnen's research strategy

Van Wijnen aims to answer two central research questions (2000: 431):

"Has candidate voting grown in importance as a result of an observed decline in cleavage voting and because parties have less distinct policy profiles?" and "Does the extent to which voters evaluate parties on the basis of their candidates' characteristics differ substantially from one party to the next?"

Immediately hereafter, he narrows 'party evaluation' (question 2) down to voting behaviour (2000: 432). So, the dependent variable in both questions is now:

voting behaviour, understood as the choice between a number of alternative electoral options. I will refer to this dependent variable as 'electoral choice'. In the Netherlands this usually boils down to the choice of a party (sometimes with a preferential vote for a candidate other than the first on the list), whereas in other systems electoral choice may primarily consist of choosing a candidate.

Van Wijnen then goes on to specify his expectations. His first expectation is:

In a depolarizing party system, voters can be expected to increasingly resort to perceived candidate characteristics as cues in deciding which of the parties deserves electoral support (2000: 435).

This expectation is based on the decreased role of traditional societal cleavages in politics.³ The second expectation is:

Parties will differ in the extent to which they are evaluated by voters on the basis of candidate evaluations (2000: 435).

This second expectation is a reformulation of the idea that there is no reason to believe a priori that all parties and candidates will be alike when it comes to the development outlined in the first expectation. Parties can be large or small, their candidates can be more or less visible in the media.

Data are available for four elections years: 1986, 1989, 1994 and 1998, and for the three or four major parties. Van Wijnen distinguishes two key indicators of orientations towards candidates: candidate evaluation (on a 100-point thermometer) and trust in prospective prime ministers (on a seven-point scale).⁴ These are the main explanatory variables. Controls are introduced for social class, church membership, church attendance, respondent's distance from the party on the left-right scale and on three structural issues (income differences, nuclear energy, and abortion/euthanasia) (2000: 436-437).

Skipping the descriptive analyses, the effects of trust in prospective prime ministers and candidate evaluations on party choice are then estimated simultaneously for the voters of PvdA, VVD, CDA and D66, for each of the four elections.

On the basis of the results, van Wijnen concludes that:

[t]rust in a candidate and sympathy for a candidate have indeed become more important for voting behaviour, both in terms of vote effects and vote explanatory power. Especially trust in the candidates of the PvdA, VVD and CDA seems to have become more important for an explanation of Dutch voting behaviour in the past two decades (2000: 448).

He also concludes that there are important differences between parties in this respect, but is unable to infer a pattern in these differences.

3 Critique of the research strategy

3.1 The causal chain

Van Wijnen's model includes two indicators for the candidate or leader factor in voting behaviour, namely the evaluation of the party leader on a 'sympathy' thermometer, and confidence in the party leader as future prime minister. He excludes the evaluation of the party on the sympathy scale. This is remarkable, because we are interested in the net impact of leader evaluations on electoral choice. In this context, 'net' refers among other things to the impact of leaders when the general evaluation of the party they lead, is taken into account. If this general evaluation of the party is *not* taken into account, the impact of leader evaluations will generally be overestimated. This follows from the theoretical causal order of the two kinds of evaluation vis-à-vis each other.

When one is interested in the causal effects of leader evaluations on vote choice, one should take other variables that might affect the relationship between leader evaluations and vote choice into account. This basic rule for assessing causality implies that third variables which precede in the theoretical causal order, those which operate at the same place in the causal chain as leader evaluations, and those which directly affect the relationship between leader evaluations and vote choice, should be included in the model. Variables that come after leader evaluations in the causal chain can be ignored when one is interested in the effect of leader evaluations on vote choice, because they do not affect the parameters of interest (i.e., the total effects of leader evaluations).

So, how are leader evaluations, party evaluations and vote choice causally related to each other? Obviously, vote choice is to be explained and comes last in the causal order. Party evaluations can only be ignored in modelling the effect of leader evaluations on vote choice if party evaluations come *after* leader evaluations in the chain. This is highly unlikely. In most democracies (though not all) political parties remain the same as their leaders come and go, to begin with. It is much more plausible to assume a two-way interaction between party and leader evaluations. But, as has just been argued, that implies that, in order to assess the effect of leader evaluations, party evaluations must be taken into account.

When this is granted, a design choice emerges. Either two correlated variables are included in a causal model, which might give rise to multicollinearity problems. Or, relevant variables are excluded from a causal model, which might lead to biased estimates. Van Wijnen opted to eliminate multicollinearity, but did not consider the possible bias in his estimates. How serious are the consequences of each of the two options?

3.2 Multicollinearity: a problem?

Van Wijnen justifies the exclusion of general party evaluation by arguing that party evaluation and leader evaluation are highly correlated (2000: 453). In any regression-type model, when two explanatory variables are highly correlated, the estimates of their effects on the dependent variable (in this case: party choice) will be relatively unreliable.

To demonstrate this, suppose that we are interested in explaining the probability of making a specific electoral choice C , $Prob(C)$. There are two important explanatory variables, the general evaluation of the party (PE) and the general evaluation of its leader (LE). A regression model may then be specified as: $Prob(C) = a + b_1PE + b_2LE + e$. In this model, a depicts the regression constant or intercept; b_1 and b_2 are the estimates of the effects of a one-unit change in PE and LE respectively on the probability of choosing C ; and e stands for the error term.

Perfect collinearity occurs in this model when PE can be written as a linear combination of LE , e.g. $PE = 3 + .75LE$. In that case, the regression model cannot be estimated since there are an infinite number of solutions for b_1 and b_2 that produce the same values of $Prob(C)$.

Perfect collinearity is a relatively rare problem. Less-than-extreme collinearity, or multicollinearity, occurs in this model when PE and LE are not linearly related, but are highly correlated in a given dataset (sample). In this case, it will be relatively difficult to separate their effects on electoral choice. This is reflected in high standard errors of the estimates of b_1 and b_2 . However, even when the intercorrelation of PE and LE is very high (but lower than 1.00), these estimates are still *unbiased*. This means that their expected value (in a hypothetical series of replications) depicts the true regression coefficient β_1 or β_2 (e.g., Berry & Feldman 1993: 190).

How serious is the impact of multicollinearity on the standard errors of the b 's, and consequently for the reliability of the results? In order to answer this question, one should first ask another: serious compared with what? Usually, the base line for comparison is the (unrealistic) situation that the correlation between regressors (PE and LE in our case) is zero. It is important to see that this base line situation is in practice about as rare as the situation of perfect collinearity, but it is a base line. In the formula for the standard error of an OLS-estimate b , the impact of other explanatory variables is reflected in a term for the variance explained when the variable is regressed on all other explanatory variables in the model.⁵ The standard error is smallest, *ceteris paribus*, when this explained variance is zero. This is what happens if, and only if, the bivariate correlation between PE and LE is zero. More generally, the smaller standard errors are obtained when the multiple

correlation coefficient between the explanatory variable of interest and all other explanatory variables is smaller.

For the OLS-model, one can accordingly calculate precisely by what factor the standard errors of estimates are inflated because of the multicollinearity in the model. This is what I have done for a model of party evaluation, leader evaluation and electoral choice.⁶ Because of data restrictions (for 1989, no party evaluations are available), only three election years – 1986, 1994, 1998 – can be analysed, in contrast with the four elections Van Wijnen considers. The results are summarized in Table 1.

Table 1 Approximate impact of collinearity on the standard errors of estimates*

1986 (N = 1,181)	PvdA	VVD	CDA	D66	
Party Evaluation	1.89	1.64	1.58	1.39	
Leader Evaluation	1.74	1.64	1.62	1.40	
1994 (N = 1,122)	PvdA	VVD	CDA	D66	GL
Party Evaluation	1.52	1.49	1.46	1.54	1.62
Leader Evaluation	1.54	1.54	1.52	1.54	1.51
1998 (N = 1,377)	PvdA	VVD	CDA	D66	GL
Party Evaluation	1.43	1.60	1.36	1.37	1.47
Leader Evaluation	1.35	1.60	1.37	1.25	1.41

* : Entries are square roots of the variance-inflation factor, $\sqrt{[1/(1-R_j^2)]}$, in which R_j^2 is the explained variance of a linear regression of the evaluation variable on all other regressors in the model.

Table 1 shows the factors by which the standard errors of the estimates of causal effects are inflated in a model that contains both party and leader evaluations as explanatory variables, compared with the hypothetical situation in which these evaluations are uncorrelated with the other explanatory variables. For example, the standard error of the estimated impact of the party evaluation of PvdA in 1986 on electoral choice is 1.89 times higher than it would have been if the party evaluation of PvdA in 1986 had been uncorrelated with the other explanatory variables, including the evaluation of PvdA-leader Den Uyl. This particular example also sets the upper limit of standard error inflation: typically, values of 1.4 - 1.6 are found.

How bad is this impact of multicollinearity? That depends, of course. Translating the results of Table 1 to Van Wijnen's results (which can only be done by way of rough approximation, since similar diagnostics are not

available for the conditional logit model), several significant coefficients for leader evaluations that he reports, will most likely become insignificant once party evaluations have been taken into account.⁷ But that is primarily due to the borderline significance of the original results in the first place.

Although no external standards are available, inflation of the standard errors by less than 1.9 does not point to serious multicollinearity.⁸ Substantively this implies that, given the importance of separating the impact of leader evaluations from the evaluations of the associated political parties, party evaluations should be included in the model.

3.3 Omitted variables

The multicollinearity problems that arise from including party evaluations in addition to leader evaluations in the same explanatory model are not very serious. What about the consequences of leaving those party evaluations out, as Van Wijnen has done?

Leaving important variables out of a regression-type model may have serious consequences for the estimated effects of those variables that are in the model. The problem is known as specification error (Berry & Feldman 1993) or omitted variable bias (King et al. 1994). To see this, consider again the simple model introduced above: $Prob(C) = a + b_1PE + b_2LE + e$. Suppose that this is the 'true' model predicting the probability of choosing C at the elections.⁹

When PE is left out of this model, it becomes part of the (new) error term $u = e + vPE$ of the model $Prob(C) = a + b_1LE + u$. Because PE is correlated with LE , this error term u will also be correlated with LE . Any regression-type model is designed to explain as much variance in the dependent variable as possible by the explanatory variables in the model. This is equivalent to minimizing the error term. So, in this example the regression mechanics will lead to an *overestimation* of the impact of LE on $Prob(C)$.

Although omitted variable bias can to some extent be regarded as the downside of multicollinearity,¹⁰ the two effects are thus very different. In the case of multicollinearity, the problem lies in the increased standard errors of the estimates, but the estimators are unbiased. In the case of omitted variables, the problem is not the standard errors of the estimates but the fact that the estimators are biased.

The problem of not including the general evaluation of the political party in a model assessing the impact of leader evaluation on electoral choice, can now be reformulated in more general terms as a choice between two options. The first is to include an explanatory variable in the model that is correlated with other explanatory variables, which leads to a decrease in the reliability of

findings. Choosing this option implies that we can be less certain that the measured impact of leader evaluations on electoral choice is not coincidental. The second option is to exclude a relevant explanatory variable from the model, which leads to a decrease in the validity of findings. Choosing this latter option (which Van Wijnen does) means that we can be seriously wrong in drawing conclusions from the measured impact of leader evaluations on electoral choice.

Somehow it has to be decided which of the two options is the worst. At this point, in linear regression models the mean square error criterion may be helpful.¹¹ However, as far as I know, there is no way yet of applying this criterion in the conditional logit model. Therefore, the only way to find out more is to test both models (including as well as excluding party evaluations), and see what comes out.

Before doing so, some other problems with Van Wijnen's analysis will be clarified.

3.4 Additional considerations

The sympathy rating of political leaders is not the only variable in Van Wijnen's model that measures leader evaluation. Another is the confidence in the leader as prospective prime minister (trust). His conclusions, quoted above, are partly based on findings for trust.

Trust and sympathy ratings are of course related to each other, not just empirically but also conceptually.¹² Moreover, since it was measured *after* the elections, the confidence rating in some party leaders as prospective prime ministers has probably been influenced by the outcome of the election. However, it can be argued that trust measures something different than leader ratings, and that is sufficient reason to include it in the replications.

Van Wijnen includes several other variables in his model. Some of these can be regarded as demographic characteristics that serve as 'standard' controls in models like his: social class and religion. The main point of including these variables is to avoid attributing explanatory power to leader evaluation data that should be attributed to these demographic factors. Other control variables in Van Wijnen's models are not so innocent. These are: party evaluations on three position issues (income differences, nuclear energy, abortion/euthanasia) and on the left-right dimension, and satisfaction with government policies. The four 'party evaluations' are measured as the absolute distance between the respondent and the party in question on each issue.

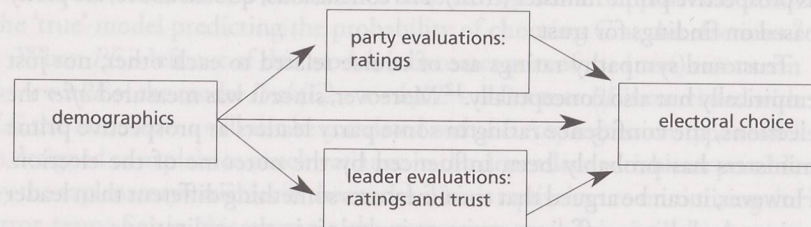
The impact of these political evaluations is not reported in the article. What is remarkable, however, is that Van Wijnen labels the distance measures as 'party evaluations'. In reality, these are typical issue orientations, which can be,

and are, used in models explaining party evaluations (or electoral choice). From other work (Aarts et al. 1999; Van Wijnen 1999), we know that (position) issue orientations are not very important (but also not unimportant) for the explanation of party evaluation and electoral choice.¹³ It seems therefore unwarranted to take issue orientations as a proxy for party evaluation. There are too many factors other than issue orientations that may determine how much people like a particular party.

Since party evaluation and leader evaluation are two general concepts, we should aim for the most general measurements that are available. These are the two questions about how sympathetic one finds the particular party of leader, on a scale from 1 to 100, supplied with the question how much confidence one has in the leader as prospective prime minister. Other, causally prior political orientations (if one wants to look at them as causally prior at all) may be disregarded.

4 New analyses

These considerations result in the following general model of the impact of leader evaluation on electoral choice:



In this model, the impact of leader evaluations should always be assessed relative to the impact of general party evaluations and controlling for the independent effect of demographic factors.

Including party evaluations has one major drawback from a viewpoint of data availability: now only the election studies in 1986, 1994 and 1998 include the data needed, since no party evaluation questions were asked in 1989. Thus, the number of time points is reduced to only three, which will make any cross-temporal conclusions even weaker than they already were in Van Wijnen's case.¹⁴

In order to assess the seriousness of omitted variable bias, I have estimated models from which leader evaluations and party evaluations have been

excluded (only the latter case will be discussed here) and contrasted these with the full model introduced above. Let me briefly rehearse the expectations underlying these analyses:

- If multicollinearity is a major problem, the standard errors of estimates of the impact of leader evaluations will be grossly inflated when party evaluations are included in the model, compared to the situation when these are excluded;
- If omitted variables cause a major bias, the estimates of the impact of leader evaluations will be considerably lower when party evaluations are included in the model, compared to the situation when these are excluded.

I have performed a series of conditional logit analyses, four for each of the three election years 1986, 1994 and 1998. The results are summarized in Tables 2–4.¹⁵

Table 2 Party evaluation, leader evaluation, and vote - 1986

A Demographic model

	PvdA	VVD	CDA
Constant	1.16 **	0.28	0
Income	-0.14 **	0.13**	0
Age	-0.02 *	-0.01	0
Gender	0.39	0.21	0
Education	-0.01	0.43**	0
Catholic	-1.95 **	-1.52**	0
Calvinist	-2.94 **	-2.34**	0
Dutch Reformed	-1.72 **	-1.14**	0

Model Characteristics

Null Model Log-likelihood	-949.20
Model Log-likelihood	-761.43
Chi-Square (df)	294.20 (14)
Adj. R ²	0.19
N =	864

B Evaluation Effects

		Party Only		Leader Only		Party and Leader	
		Effect	s.e.	Effect	s.e.	Effect	s.e.
PvdA	Party Rating	10.13**	0.94	-	-	10.10**	1.26
	Leader Rating	-	-	3.94**	0.64	-1.67	1.11
	Trust	-	-	0.69**	0.08	0.55**	0.11
VVD	Party rating	11.60**	1.00	-	-	11.48**	1.20
	Leader Rating	-	-	5.56**	0.69	-0.44	0.97
	Trust	-	-	0.29**	0.08	0.24	0.10
CDA	Party Rating	7.05**	0.74	-	-	4.71**	0.83
	Leader Rating	-	-	4.47**	0.70	2.25	0.98
	Trust	-	-	0.27 *	0.08	0.26 *	0.10

Model Characteristics

Null Model Log-likelihood	-949.20	-949.20	-949.20
Model Log-likelihood	-304.53	-425.43	-275.81
Chi-Square (df)	1208.01 (17)	966.19 (20)	1265.44 (23)
Adj. R ²	0.68	0.55	0.71
N =	864	864	864
Perc. Correctly Predicted	81%	71%	82%

*: p < 0.01; **: p < 0.001, one-tailed test

Notes: Estimates of conditional logit model obtained using the programme LIMDEP 7.0.
 Dependent variables are the choice sets {PvdA, VVD, CDA} (1986) or {PvdA, VVD, CDA, D66} (1994, 1998). Data are not weighted.

Table 3 Party evaluation, leader evaluation, and vote - 1994

A Demographic model

	PvdA	VVD	CDA	D66
Constant	0.24	-0.19	-1.80**	0
Income	-0.20	0.89 *	0.46	0
Age	0.04 **	0.03 **	0.05**	0
Gender	0.20	0.52 *	0.36	0
Education	0.02	0.12	0.17	0
Catholic	-0.48	0.37	2.57**	0
Calvinist	-2.45	0.15	3.18**	0
Dutch Reformed	0.33	0.97 *	2.45**	0

Model Characteristics

Null Model Log-likelihood	-1336.39
Model Log-likelihood	-1144.33
Chi-Square (df)	370.65 (21)
Adj. R ²	0.14
N =	964

B Evaluation effects

		Party Only		Leader Only		Party and Leader	
		Effect	s.e.	Effect	s.e.	Effect	s.e.
PvdA	Party Rating	11.68**	0.81	-	-	9.60**	1.00
	Leader Rating	-	-	7.19**	0.66	2.60 *	0.94
	Trust	-	-	0.37**	0.08	0.26 *	0.09
VVD	Party rating	12.81**	0.88	-	-	11.05**	1.00
	Leader Rating	-	-	8.26**	0.60	2.65 *	0.82
	Trust	-	-	-0.14	0.07	-0.02	0.09
CDA	Party Rating	12.59**	0.93	-	-	11.22**	1.06
	Leader Rating	-	-	7.15**	0.62	1.98	0.86
	Trust	-	-	-0.29**	0.07	-0.33**	0.08
D66	Party Rating	11.10**	0.84	-	-	10.11**	1.07
	Leader Rating	-	-	7.90**	0.66	1.71	0.93
	Trust	-	-	-0.02	0.08	-0.06	0.10

Model Characteristics

Null Model Log-likelihood	-1336.36	-1336.36	-1336.36
Model Log-likelihood	-560.66	-770.56	-534.91
Chi-Square (df)	1537.99 (25)	1118.20 (29)	1589.50(33)
Adj. R ²	0.58	0.42	0.59
N =	964	964	964
Perc. Correctly Predicted	69%	56%	70%

*: p < 0.01; **: p < 0.001, one-tailed test

Notes: Estimates of conditional logit model obtained using the programme LIMDEP 7.0.
 Dependent variables are the choice sets {PvdA, VVD, CDA} (1986) or {PvdA, VVD, CDA, D66} (1994, 1998). Data are not weighted

Table 4 Party evaluation, leader evaluation, and vote - 1998

A Demographic model

	PvdA	VVD	CDA	D66
Constant	1.26 **	0.32	-1.30**	0
Income	-0.06	0.10 *	-0.05	0
Age	0.01	0.00	0.03**	0
Gender	0.15	0.89 **	0.71 *	0
Education	-0.64 **	-0.60 **	-0.59**	0
Catholic	0.13	0.37	2.46**	0
Calvinist	0.55	1.60	4.45**	0
Dutch Reformed	0.52	1.05	2.89**	0

Model Characteristics

Null Model Log-likelihood	-1369.66
Model Log-likelihood	-1140.99
Chi-Square (df)	386.46 (21)
Adj. R ²	0.16
N =	988

B Evaluation effects

		Party Only		Leader Only		Party and Leader	
		Effect	s.e.	Effect	s.e.	Effect	s.e.
PvdA	Party Rating	12.07 **	0.83	-	-	10.92**	1.01
	Leader Rating	-	-	3.41**	0.70	-0.29	0.94
	Trust	-	-	0.92**	0.14	0.91**	0.17
VVD	Party rating	11.25 **	0.80	-	-	8.10**	1.05
	Leader Rating	-	-	4.22**	0.63	1.06	0.92
	Trust	-	-	0.95**	0.09	0.87**	0.11
CDA	Party Rating	11.79 **	0.88	-	-	8.84**	1.09
	Leader Rating	-	-	6.25**	0.80	2.81 *	0.97
	Trust	-	-	0.75**	0.10	0.65**	0.11
D66	Party Rating	12.65 **	0.98	-	-	11.56**	1.09
	Leader Rating	-	-	3.34**	0.68	0.62	0.84
	Trust	-	-	0.46**	0.08	0.41**	0.10

Model Characteristics

Null Model Log-likelihood	-1369.66	-1369.66	-1369.66
Model Log-likelihood	-624.94	-745.06	-534.08
Chi-Square (df)	1418.58 (25)	1178.32 (29)	1600.28 (33)
Adj. R ²	0.54	0.45	0.61
N =	988	988	988
Perc. Correctly Predicted	66%	58%	70%

*: p < 0.01; **: p < 0.001, one-tailed test

Notes: Estimates of conditional logit model obtained using the programme LIMDEP 7.0.

Dependent variables are the choice sets {PvdA, VVD, CDA} (1986) or {PvdA, VVD, CDA, D66} (1994, 1998). Data are not weighted

The tables consist of two parts. First, the results for a strictly demographic multinomial logit model are presented. This model is interesting for two reasons. It shows how, in different elections, electoral choice (limited to the few parties in the analysis) depends on people's social background, and it sets a baseline for explained variance. For reasons of identification, the demographic effects for one of the parties in the choice set have to be set to zero; this is shown in the tables. Secondly, the results for conditional logit analyses of party and leader evaluation on top of the demographic models are presented. For each election, three conditional logit analyses have been performed: one with party evaluations only, one with leader evaluations (ratings and trust) only, and one with both party and leader evaluations.

In the multinomial logit model, the explanatory variables are regarded as individual-specific. This means that, for a given individual, the *values* these variables take are the same for all choice options, whereas in a group of individuals their *effect* may vary over these options. For example, a voter has a certain age, but the effect of age on vote choice may vary from party to party. In contrast, the conditional logit model regards the variables as attributes of the choice options. This means that, for a given individual, their *values* may vary over the choice options, but for a group their *effect* on the parties is constrained to be equal. For example, a party is characterized by a certain level of trust in its leader, and this trust in a specific leader has the same effect on all parties in the choice set.¹⁶ But a voter may trust the leader of one party more than that of another.

The demographic models (Parts A of the tables) show a familiar pattern of effects for the different parties. For example, in 1986 it is seen that in contrast with CDA, the probability of voting PvdA decreases as household income increases, whereas the relationship is positive in the case of VVD. For 1994 and 1998, the contrasts are shown with D66 as reference party. It is seen that in 1994 D66 is especially attractive for younger people; in 1998, with the number of D66 voters drastically reduced, the age factor is absent in the competition with the 'purple coalition' partners PvdA and VVD. Religion is

and remains very important for the CDA-vote. The explained variance in party choice as a result of the demographic variables ranges from 19% in 1986, via 14% in 1994, to 16% in 1998.¹⁷

The evaluation models are presented in Parts B of the tables. Note that these models also include the demographic variables, but for reasons of space their effects are not shown. The most obvious finding is the huge difference in the impact of leader evaluations in a model that excludes party evaluations (Van Wijnen's strategy), compared with the model that includes party evaluations (my strategy). In the former case, all eleven estimated effects of sympathy ratings, and nine of the eleven effects of trust in leaders, are clearly present (as they are in the comparable analyses of Van Wijnen). In the latter case, eight of the eleven effects of sympathy ratings do not even reach statistical significance, and the other three are considerably lower than those in the 'Leader Only' columns of the tables. The effects of trust also decrease once party evaluations are taken into account, but now all but one (i.e., eight out of nine) remain statistically significant, and the decrease is in some cases hardly noticeable. Especially the findings for sympathy ratings, but also those for trust, underline the importance of omitted variable bias. Because of the correlation of leader evaluations with other explanatory variables, notably the corresponding party evaluation, exclusion of relevant explanatory variables (party evaluation) distorts the conclusions that one might be tempted to draw about the impact of leaders on electoral choice. Expectation (2) is borne out by the data. It can also be seen that the increase in the standard errors of the estimates in the 'Party and Leader' columns when compared with the 'Leader Only' columns, is generally even lower than what was expected on the basis of the (strictly speaking inappropriate) variance inflation factors in Table 1. To be sure, the standard errors are inflated, and in this sense expectation (1) is also borne out. However, as argued above, the increase of the standard errors is not extreme. The same cannot be said for the bias in the estimates resulting from the exclusion of an important variable.

A final observation on Tables 2-4 focuses on the explanatory power of the models. By taking the adjusted R^2 -measures, it can be seen that, once party evaluations are being considered, the evaluations of leaders do not add much explanatory power to the model. This finding emphasizes the party-centered character of electoral choice in the Netherlands. Leaders sometimes matter, to be sure (whether in a positive or in a negative way – the model does not permit conclusions about direction), but when it comes to electoral choice they are overshadowed by the orientations towards their parties.

How do these findings bear on Van Wijnen's two central results, referred to earlier? Recall that the first of his results is that trust in a candidate and sympathy for a candidate have become more important for voting behaviour, and that the second pertains to (unsystematic) differences between parties.

Consider again Parts B of Tables 2-4. For the three major parties, PvdA, VVD and CDA, the coefficients in the 'Party and Leader' columns for neither leader ratings nor trust show a 'continuous' increase (over three elections!). This is insufficient groundwork for a statement on growing leadership effects. The second conclusion therefore is not based on evidence either.

The message of the election-specific information in Tables 2-4 is visualized in the three panels of Figure 1, each of which shows the development through time of the effect of leader rating on electoral choice, both in the Leader Only-model and in the Party and Leader-model, for PvdA, VVD, and CDA. The second line is the more interesting since it shows the net effect of leader ratings. Also, confidence intervals for the estimated effects are shown, based on $b \pm 2 \times \text{standard error of } b$.

Figure 1 Impact of PVDA leader on electoral choice

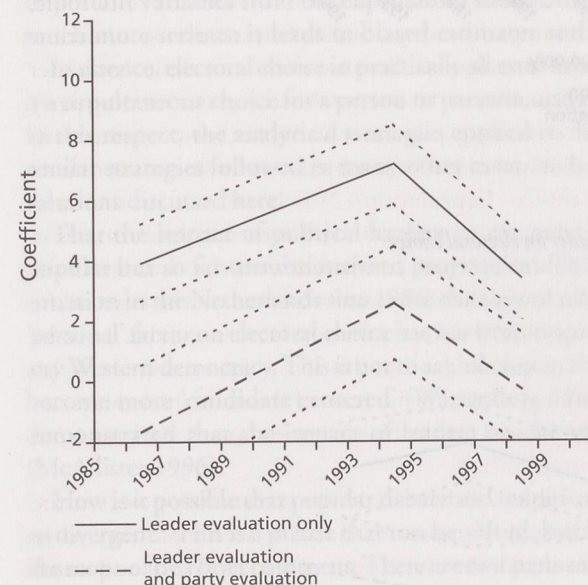
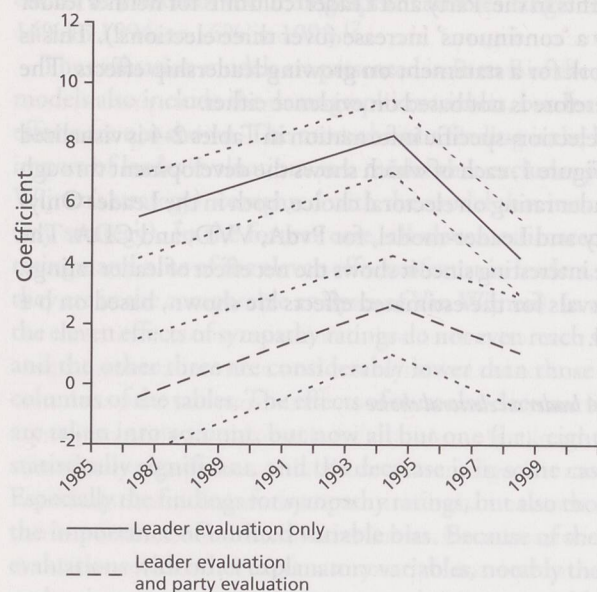
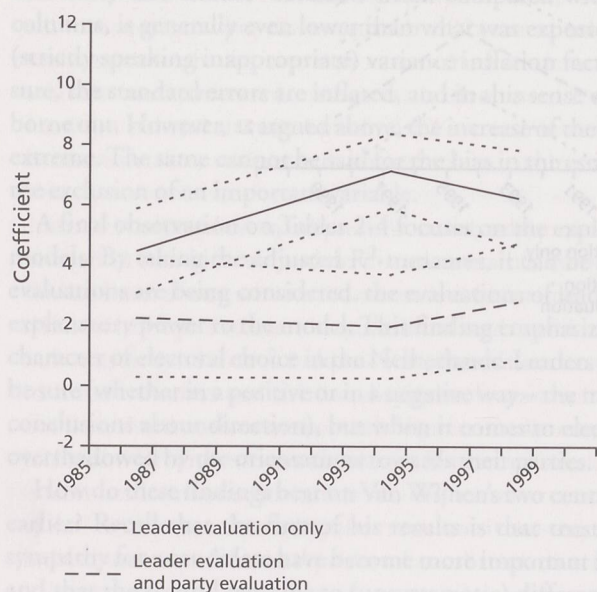


Figure 2 *Impact of VVD leader on electoral choice*Figure 3 *Impact of CDA leader on electoral choice*

Three points in time are definitely insufficient to infer any general trend, even for the short period (1986–1998) being considered here. But if one still wants to draw conclusions from the panels in Figure 1, the obvious conclusion seems to be (1) that independent effects of leader ratings are usually absent (the value 0 falls within the confidence interval), and (2) that where leader rating effects are present, they do not reveal an overall pattern. It should be added that similar graphs for the effects of trust would show less impressive differences, as noted before.

5 Discussion

The interwovenness of political parties and their leaders makes it difficult to investigate the 'net' effects of leader evaluations on electoral choice. I have argued, however, that it would be wrong to shy back from assessing such net effects out of fear for unreliable results. The downside of this, omitting important variables from the explanatory model, has consequences that are much more serious: it leads to biased estimates and mistaken conclusions.

In essence, electoral choice in practically all established democratic countries is a simultaneous choice for a person or persons, and for a party organization. In this respect, the analytical strategies applied to the Dutch case stand for similar strategies followed in many other cases, including the problems and solutions discussed here.

That the impact of political leaders on electoral choice is growing, is a popular but so far unsubstantiated proposition. This is not unique for the situation in the Netherlands since 1986: the supposedly growing impact of the 'personal' factor on electoral choice has not been empirically demonstrated for any Western democracy. This is not to say, of course, that politics may not have become more 'candidate centered' (Wattenberg 1991). It just has not been demonstrated that the impact of leaders on the voting act has increased (McAllister 1996).

How is it possible that popular debate and empirical evidence have become so divergent? This is a puzzle that can be solved, but it would take us beyond the scope of this brief comment. There are two parts of the puzzle that I submit without detailing them. First, there has been a shortage of empirical research of sufficient quality to counterbalance the simple but attractive notion that TV can be blamed. The rise of televised politics has certainly been an important factor in the transformation of the modern society, but it is all too easy to suggest that television has led to a focus on the persons instead of organizations, programmes or issues. For one thing, the amount of candidate-centered politics of the pre-television age tends to be underestimated. Public rallies, newspapers and radio broadcasts have emphasized the persons behind the parties for a long time.

Secondly, the problem is simply difficult to investigate. The dependent variable is a choice set (in our analyses: four or five larger parties; more generally: electoral choice including the option of not voting). This means that all kinds of traditional (statistical) models are less appropriate. One of the merits of Van Wijnen's contribution compared with earlier work on the Dutch case is, as mentioned earlier, that he uses a model specifically designed to deal with a nominal choice variable. But even in the conditional logit model, it is not easy to interpret the results. To give one example: it is commonly understood that the incumbent prime minister, Wim Kok, formed an important factor in the electoral success of his party, PvdA, in 1998. We therefore might expect to find a leader rating effect on electoral choice in this case, besides the effect of trust in Kok as a prime minister. Table 4 shows, however, that sympathy for Kok has no significant impact on electoral choice over and above the evaluation of his party. How should this counterintuitive finding be interpreted?

The clue to understanding lies in the realization that a political leader who is acclaimed by (almost) everyone, will get a high average evaluation score and little variation around this average. But, if there is little variation in the evaluation, the evaluation cannot contribute much to an understanding of variation in electoral choice – which is what the regression model aims at. Counterfactual reasoning might help out here: what would have happened to the vote share of the PvdA if Kok had been evaluated on average 20 points lower and the variance of his evaluation had been twice as high? This is exactly what Van Wijnen tries to do in his contribution (2000: 446). But the validity of counterfactual reasoning hinges on the quality of the explanatory model – and I have argued that this quality can be questioned.

Notes

1. Some of the arguments in this paper and part of the empirical analyses have been taken from an earlier conference paper (Aarts 2000), and have been developed in the context of an international research project, 'The Changing Impact of Leaders'. For specific suggestions, I thank André Blais, Henk van der Kolk, and anonymous referees of this journal.

The data for the Dutch national election studies of 1986, 1994, and 1998 are available through NIWI-Steinmetz Archive in Amsterdam and other major data archives. All syntax files used for the analyses in this paper can be obtained from the author (e-mail c.w.a.m.aarts@bsk.utwente.nl).

2. Of course, countless analyses have been published on the impact of candidates or political leaders on voting behaviour, using the full range of methods from simple descriptives to complicated structural models. In this comment I will, with some

exceptions, not cite this large body of literature, as my goal is more specific. References to analyses for the Netherlands can be found in Van Wijnen (2000) or in Aarts (2000).

3. There is an implicit additional assumption underlying the first expectation: voters are always subject to a system of political push and pull factors, and if one factor becomes less important, other factors will accordingly increase in significance. This reflects a fundamental *horror vacui* in electoral research. But even though it is a very common assumption, it is not necessarily true. It is conceivable that with the decrease of traditional political cleavages, the random component of behaviour increases. Unless one believes in a kind of social determinism, the possibility that our models perform ever worse should be faced. I will, however, not go into evidence for or against increasing randomness here.

4. In his analyses, the thermometer scores have been rescaled to the 1-7 interval, so that coefficients for the two key explanatory variables are directly comparable.

5. The formula is: standard error of $b_i = s_{b_i} = \sqrt{\{\sum_j (Y_j - Y_i)^2 / \sum_j (X_{ij} - X_{\text{mean}})^2 (1 - R_i^2) (n - k - 1)\}}$ (Berry & Feldman 1993: 163), in which R_i^2 is the variance explained by a regression of X_i on all other X-variables.

6. The model contains the following variables: party evaluations of PvdA, VVD, CDA, D66, and (only in 1994 and 1998) GroenLinks; leader evaluations of the leaders of PvdA, VVD, CDA, D66, and (again only in 1994 and 1998) GroenLinks; gender, age, household income, and educational level. For each regression, one of the evaluations was taken as dependent variable, and all others as explanatory variables. The inclusion of all other evaluations in each regression equation most closely resembles the conditional logit model to be estimated later.

The names of the party leaders, and the bivariate correlations between the party and leader evaluations are as follows:

	1986	r	1994	r	1998	r
PvdA	den Uyl	0.80	Kok	0.70	Kok	0.62
VVD	Nijpels	0.73	Bolkestein	0.69	Bolkestein	0.70
CDA	Lubbers	0.72	Brinkman	0.67	De Hoop Scheffer	0.64
D66	van Mierlo	0.65	van Mierlo	0.71	Borst	0.51
GL			Brouwer	0.70	Rosenmöller	0.67

7. Table 6 in Van Wijnen (2000). Note that Van Wijnen rescaled the leader evaluation variables to the 1-7 range, which makes his coefficients not directly comparable to those reported in this comment.

8. Cf. Fox 1993: 253.

9. 'True' in this context means that (a) the structure of the model is correct, and (b) there are no important other variables that contribute to the explanation of Prob(C). In everyday research practice, it means that there is no better model available.

10. In both cases, the problems disappear when there is no correlation between the regressors. In the case of omitted variables, bias will also be absent when the omitted variable is uncorrelated with the dependent variable.

11. The Mean Squared Error is defined for an estimator b of 'true target' β as $E(b - \beta)^2$, which can be reformulated as the sum of variance and bias: $MSE = var(b) + (\mu_b - \beta)^2$, in which $\mu_b = E(b)$; refer to Wonnacott & Wonnacott 1979: 61.

12. The bivariate correlations coefficients between trust and leader evaluation are as follows:

	1986	1994	1998
PvdA	0.64	0.45	0.54
VVD	0.52	0.46	0.64
CDA	0.66	0.57	0.53
D66	—	0.42	0.48

13. Aarts, Macdonald & Rabinowitz (1999) report additional explained variances in party evaluations as a result of position issues and general left-right ideology (over a simple demographic explanatory model) in the range of 0-34%. The variation over elections, parties, and over issue evaluation models is large. Left-right ideology is consistently the single most important contributor to the explained variances. Comparable unpublished analyses of the 1998 elections by the same authors confirm these findings.

14. Including trust next to sympathy ratings also reduces the number of observations: for 1986, no trust data are available for the leader of D66, and in 1994 and 1998 the trust question was not asked for the leader of Green Left.

15. The respondents included in the analyses are those with valid scores on all variables, and who voted for either PvdA, VVD, CDA, or D66 (in 1986, D66 voters were not included; see previous note). The demographic variables are slightly different from those in Van Wijnen (2000), but this is of little consequence for the results. Note that the sympathy ratings for both parties and leaders have been rescaled to the 0-1 interval in order to avoid clumsily small numbers in the tables. If one wants to compare the coefficients for the ratings with those for trust, the coefficients for trust should be multiplied by 7.

16. The conditional logit model can be taken further by assuming that the 'proper names' of the parties and/or leaders in the evaluation data can be missed. This is a theoretical decision; there are no problems in handling such even more abstract models statistically. In my opinion, by following this strategy one would miss the point of analysing leader effects. This is, however, a point that cannot be elaborated in the context of this note.

17. It should be noted that the relatively high figure for 1986 may partly result from the exclusion of D66 (a party with a relatively weak demographic profile) from the analysis.

Bibliography

- Aarts, Kees, Stuart E. Macdonald and George Rabinowitz (1999), 'Issues and party competition in the Netherlands', *Comparative Political Studies* 32(1), pp. 63-99.
- Aarts, Kees (2000), 'The Impact of Leader Evaluations on Voting in The Netherlands', paper presented at the Annual Meeting of the American Political Science Association, Washington DC, August 31 - September 3.
- Berry, William D., and Stanley Feldman (1993), 'Multiple regression in practice', in: Michael S. Lewis-Beck (ed.), *Regression Analysis*. International Handbooks of Quantitative Applications in the Social Sciences, Volume 2. London: Sage.
- Fox, John (1993), 'Regression diagnostics', in: Michael S. Lewis-Beck (ed.), *Regression Analysis*. International Handbooks of Quantitative Applications in the Social Sciences, Volume 2. London: Sage.
- King, Gary, Robert O. Keohane and Sidney Verba (1994). *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton: Princeton University Press.
- Lewis-Beck, Michael S. (ed.) (1993), *Regression Analysis*. International Handbooks of Quantitative Applications in the Social Sciences, Volume 2. London: Sage.
- McAllister, Ian (1996), 'Leaders', in: Lawrence LeDuc, Richard G. Niemi and Pippa Norris (eds.), *Comparing Democracies: Elections and Voting in Global Perspective*. Thousand Oaks: Sage.
- Van Wijnen, Pieter (1999), 'The role of policy preferences in the Dutch national elections of 1998', *Acta Politica* 34 (Summer/Autumn), pp. 200-235.
- Van Wijnen, Pieter (2000), 'Candidates and voting behaviour', *Acta Politica* 35 (Winter), pp. 430-458.
- Wattenberg, Martin P. (1991), *The Rise of Candidate-Centered Politics: Presidential Elections of the 1980s*. Cambridge, Mass.: Harvard University Press.
- Wonnacott, Ronald J., and Thomas H. Wonnacott (1979), *Econometrics*. 2nd Edition. New York: Wiley.